

142

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: Sheppard, Paul O.
- (ii) TITLE OF THE INVENTION: SERINE PROTEASE POLYPEPTIDES  
AND MATERIALS AND METHODS FOR MAKING THEM
- (iii) NUMBER OF SEQUENCES: 18
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: ZymoGenetics, Inc.  
(B) STREET: 1201 Eastlake Avenue East  
(C) CITY: Seattle  
(D) STATE: WA  
(E) COUNTRY: USA  
(F) ZIP: 98102
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Diskette  
(B) COMPUTER: IBM Compatible  
(C) OPERATING SYSTEM: DOS  
(D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER:  
(B) FILING DATE:
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: Parker, Gary E  
(B) REGISTRATION NUMBER: 31,648  
(C) REFERENCE/DOCKET NUMBER: 97-16C1
- (ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: 206-442-6673  
(B) TELEFAX: 206-442-6678  
(C) TELEX:

43

008060" 22985960

10430

2 43

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1634 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 105...1280
- (D) OTHER INFORMATION:

- (A) NAME/KEY: Signal Sequence
- (B) LOCATION: 105...161
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGCACGAGGG	GGAGCCGCGC	GCTCTCTCCC	GGCGCCACAC	CCTGTCTGAG	CGGCGCAGCG	60
AGCCGCGGCC	CGGGCGGGCT	GCTCGGCGCG	GAACAGTGCT	CGGC ATG GCA GGG ATT		116
				Met Ala Gly Ile		
CCA GGG CTC CTC TTC CTT CTC TTC TTT CTG CTC TGT GCT GTT GGG CAA						164
Pro Gly Leu Leu Phe Leu Leu Phe Phe Leu Leu Cys Ala Val Gly Gln						
-15	-10	-5		1		
GTG AGC CCT TAC AGT GCC CCC TGG AAA CCC ACT TGG CCT GCA TAC CGC						212
Val Ser Pro Tyr Ser Ala Pro Trp Lys Pro Thr Trp Pro Ala Tyr Arg						
5	10	15				
CTC CCT GTC GTC TTG CCC CAG TCT ACC CTC AAT TTA GCC AAG CCA GAC						260
Leu Pro Val Val Leu Pro Gln Ser Thr Leu Asn Leu Ala Lys Pro Asp						
20	25	30				
TTT GGA GCC GAA GCC AAA TTA GAA GTA TCT TCT TCA TGT GGA CCC CAG						308
Phe Gly Ala Glu Ala Lys Leu Glu Val Ser Ser Ser Cys Gly Pro Gln						
35	40	45				

44

008060 2299950

244

TGT CAT AAG GGA ACT CCA CTG CCC ACT TAC AAA GAA GCC AAG CAA TAT	356
Cys His Lys Gly Thr Pro Leu Pro Thr Tyr Lys Glu Ala Lys Gln Tyr	
50 55 60 65	
CTG TCT TAT GAA ACG CTC TAT GCC AAT GGC AGC CGC ACA GAG ACN CAG	404
Leu Ser Tyr Glu Thr Leu Tyr Ala Asn Gly Ser Arg Thr Glu Xaa Gln	
70 75 80	
GTG GGC ATC TAC ATC CTC AGC AGT AGT GGA GAT GGG GCC CAN CNC CGA	452
Val Gly Ile Tyr Ile Leu Ser Ser Ser Gly Asp Gly Ala Xaa Xaa Arg	
85 90 95	
GAC TCA GGG TCT TCA GGA AAG TCT CGA AGG AAG CGG CAG ATT TAT GGC	500
Asp Ser Gly Ser Ser Gly Lys Ser Arg Arg Lys Arg Gln Ile Tyr Gly	
100 105 110	
TAT GAC AGC AGG TTC AGC ATT TTT GGG AAG GAC TTC CTG CTC AAC TAC	548
Tyr Asp Ser Arg Phe Ser Ile Phe Gly Lys Asp Phe Leu Leu Asn Tyr	
115 120 125	
CCT TTC TCA ACA TCA GTG AAG TTA TCC ACG GGC TGC ACC GGC ACC CTG	596
Pro Phe Ser Thr Ser Val Lys Leu Ser Thr Gly Cys Thr Gly Thr Leu	
130 135 140 145	
GTG GCA GAA AAN CAT GTC CTC ACA GCT GCC CAC TGC ATA CAC GAT GGA	644
Val Ala Glu Xaa His Val Leu Thr Ala Ala His Cys Ile His Asp Gly	
150 155 160	
AAA ACC TAT GTG AAA GGA ACC CAG AAG CTT CGA GTC GGC TTC CTA AAG	692
Lys Thr Tyr Val Lys Gly Thr Gln Lys Leu Arg Val Gly Phe Leu Lys	
165 170 175	
CCC AAG TTT AAA GAT GGT GGT CGA GGG GCC AAC GAC TCC ACT TCA GCC	740
Pro Lys Phe Lys Asp Gly Gly Arg Gly Ala Asn Asp Ser Thr Ser Ala	
180 185 190	
ATG CCC GAG CAG ATG AAA TTT CAG TGG ATC CGG GTG AAA CGC ACC CAT	788
Met Pro Glu Gln Met Lys Phe Gln Trp Ile Arg Val Lys Arg Thr His	
195 200 205	
GTG CCC AAG GGT TGG ATC AAG GGC AAT GCC AAT GAC ATC GGC ATG GAT	836
Val Pro Lys Gly Trp Ile Lys Gly Asn Ala Asn Asp Ile Gly Met Asp	
210 215 220 225	

008060" 22985960

45

145

TAT GAT TAT GCC CTC CTG GAA CTC AAA AAG CCC CAC AAG AGA AAA TTT 884  
Tyr Asp Tyr Ala Leu Leu Glu Leu Lys Lys Pro His Lys Arg Lys Phe  
230 235 240

ATG AAG ATT GGG GTG AGC CCT CCT GCT AAG CAG CTG CCA GGG GGC AGA 932  
Met Lys Ile Gly Val Ser Pro Pro Ala Lys Gln Leu Pro Gly Gly Arg  
245 250 255

ATT CAC TTC TCT GGT TAT GAC AAT GAC CGA CCA GGC AAT TTG GTG TAT 980  
Ile His Phe Ser Gly Tyr Asp Asn Asp Arg Pro Gly Asn Leu Val Tyr  
260 265 270

CGC TTC TGT GAC GTC AAA GAC GAG ACC TAT GAC TTG TTG TAC CAG CAA 1028  
Arg Phe Cys Asp Val Lys Asp Glu Thr Tyr Asp Leu Leu Tyr Gln Gln  
275 280 285

TGC GAT GCC CAG CCA GGG GCC AGC GGG TAT GGG GTA TAT GTG AGG ATG 1076  
Cys Asp Ala Gln Pro Gly Ala Ser Gly Tyr Gly Val Tyr Val Arg Met  
290 295 300 305

TGG AAG AGA CAG CAG CAG AAG TGG GAG CGA AAA ATT ATT GGC ATT TTT 1124  
Trp Lys Arg Gln Gln Gln Lys Trp Glu Arg Lys Ile Ile Gly Ile Phe  
310 315 320

TCA GGG CAC CAG TGG GTG GAC ATG AAT GGT TCC CCA CAG GAT TTC AAC 1172  
Ser Gly His Gln Trp Val Asp Met Asn Gly Ser Pro Gln Asp Phe Asn  
325 330 335

GTG GCT GTC AGA ATC ACT CCT CTC AAA TAT GCC CAG ATC TGC TAT TGG 1220  
Val Ala Val Arg Ile Thr Pro Leu Lys Tyr Ala Gln Ile Cys Tyr Trp  
340 345 350

ATT AAA GGA AAC TAC CTG GAT TGT AGG GAG GGT GAC ACA GTG TTC CTT 1268  
Ile Lys Gly Asn Tyr Leu Asp Cys Arg Glu Gly Asp Thr Val Phe Leu  
355 360 365

CCT GGC AGC AAT TAAGGTCTTC ATGTTCTTAT TTTAGGAGAG GCCAAATTGT TTTT 1325  
Pro Gly Ser Asn  
370

GTCATTGGCG TGCACACGTG TGTGTGTGTG TGTGTGTGTG TGTAAGGTGT CTTATAATCT 1385  
TTTACCTATT TCTTACAATT GCAAGATGAC TGGCTTTACT ATTTGAAAAC TGGTTTGTGT 1445  
ATCATATCAT ATATCATTTA AGCAGTTTGA AGGCATACTT TTGCATAGAA ATAAAAAAAA 1505  
TACTGATTTG GGGCAATGAG GAATATTTGA CAATTAAGTT AATCTTCACG TTTTTCGAAA 1565  
CTTTGATTTT TATTTTCATCT GAACTTGTTT CAAAGATTTA TATTAAATAT TTGGCATACA 1625

46

008060" 22985960



847

Gly Phe Leu Lys Pro Lys Phe Lys Asp Gly Gly Arg Gly Ala Asn Asp  
 175 180 185  
 Ser Thr Ser Ala Met Pro Glu Gln Met Lys Phe Gln Trp Ile Arg Val  
 190 195 200 205  
 Lys Arg Thr His Val Pro Lys Gly Trp Ile Lys Gly Asn Ala Asn Asp  
 210 215 220  
 Ile Gly Met Asp Tyr Asp Tyr Ala Leu Leu Glu Leu Lys Lys Pro His  
 225 230 235  
 Lys Arg Lys Phe Met Lys Ile Gly Val Ser Pro Pro Ala Lys Gln Leu  
 240 245 250  
 Pro Gly Gly Arg Ile His Phe Ser Gly Tyr Asp Asn Asp Arg Pro Gly  
 255 260 265  
 Asn Leu Val Tyr Arg Phe Cys Asp Val Lys Asp Glu Thr Tyr Asp Leu  
 270 275 280 285  
 Leu Tyr Gln Gln Cys Asp Ala Gln Pro Gly Ala Ser Gly Tyr Gly Val  
 290 295 300  
 Tyr Val Arg Met Trp Lys Arg Gln Gln Gln Lys Trp Glu Arg Lys Ile  
 305 310 315  
 Ile Gly Ile Phe Ser Gly His Gln Trp Val Asp Met Asn Gly Ser Pro  
 320 325 330  
 Gln Asp Phe Asn Val Ala Val Arg Ile Thr Pro Leu Lys Tyr Ala Gln  
 335 340 345  
 Ile Cys Tyr Trp Ile Lys Gly Asn Tyr Leu Asp Cys Arg Glu Gly Asp  
 350 355 360 365  
 Thr Val Phe Leu Pro Gly Ser Asn  
 370

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TGYACNGGNW SNHTNRT

17

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid

48

000000" 4988960

148

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AYNADNSWNC CNGTRCA

17

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ACNGCNGSNC AYTGYAT

17

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATRCARTGNS CNGCNGT

17

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

49

008060 44985960

8/49

WYRTNCCNWV NGGNTGG

17

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CCANCCNBWN GGNAYRW

17

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AYNRAYTAYG AYTAYGS

17

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

SCRTARTCRT ARTYNRT

17

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs

008060" 22985960

50



50

[illegible]

TATGCAGGCC AAGTGGGTTT CCAGGGGGCA CTGTAAGGGC

40

(ii) MOLECULE TYPE: Other  
(vii) IMMEDIATE SOURCE:  
(B) CLONE: ZC13508

TCTGCTCTGT GCTGTTGG

18

(ii) MOLECULE TYPE: Other  
(vii) IMMEDIATE SOURCE:  
(B) CLONE: ZC13509

AGTCTGGCTT GGCTAAAT

18

51

1051

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1656 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 105...1280
- (D) OTHER INFORMATION:

- (A) NAME/KEY: Signal Sequence
- (B) LOCATION: 105...161
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GGC	CAC	GAG	GGG	GG	AGC	CGC	GC	GCT	CT	CT	CCC	GG	CGC	CCC	CAC	CCT	GT	CT	GAG	CG	GC	CAG	CG		60
AGC	CGC	GGC	CGG	CGG	GCT	GCT	CGC	GC	GA	CAG	TG	CT	CGG	ATG	GCA	GGG	ATT								116
																									Met Ala Gly Ile
CCA	GGG	CTC	CTC	TTC	CTT	CTC	TTC	TTT	CTG	CTC	TGT	GCT	GTT	GGG	CAA										164
Pro	Gly	Leu	Leu	Phe	Leu	Leu	Phe	Phe	Leu	Leu	Cys	Ala	Val	Gly	Gln										
-15																									1
GTG	AGC	CCT	TAC	AGT	GCC	CCC	TGG	AAA	CCC	ACT	TGG	CCT	GCA	TAC	CGC										212
Val	Ser	Pro	Tyr	Ser	Ala	Pro	Trp	Lys	Pro	Thr	Trp	Pro	Ala	Tyr	Arg										
CTC	CCT	GTC	GTC	TTG	CCC	CAG	TCT	ACC	CTC	AAT	TTA	GCC	AAG	CCA	GAC										260
Leu	Pro	Val	Val	Leu	Pro	Gln	Ser	Thr	Leu	Asn	Leu	Ala	Lys	Pro	Asp										
TTT	GGA	GCC	GAA	GCC	AAA	TTA	GAA	GTA	TCT	TCT	TCA	TGT	GGA	CCC	CAG										308
Phe	Gly	Ala	Glu	Ala	Lys	Leu	Glu	Val	Ser	Ser	Ser	Cys	Gly	Pro	Gln										
TGT	CAT	AAG	GGA	ACT	CCA	CTG	CCC	ACT	TAC	GAA	GAG	GCC	AAG	CAA	TAT										356
Cys	His	Lys	Gly	Thr	Pro	Leu	Pro	Thr	Tyr	Glu	Glu	Ala	Lys	Gln	Tyr										
50																									

52

008060" 2985960

152

CTG TCT TAT GAA ACG CTC TAT GCC AAT GGC AGC CGC ACA GAG ACG CAG	404
Leu Ser Tyr Glu Thr Leu Tyr Ala Asn Gly Ser Arg Thr Glu Thr Gln	
70 75 80	
GTG GGC ATC TAC ATC CTC AGC AGT AGT GGA GAT GGG GCC CAA CAC CGA	452
Val Gly Ile Tyr Ile Leu Ser Ser Ser Gly Asp Gly Ala Gln His Arg	
85 90 95	
GAC TCA GGG TCT TCA GGA AAG TCT CGA AGG AAG CGG CAG ATT TAT GGC	500
Asp Ser Gly Ser Ser Gly Lys Ser Arg Arg Lys Arg Gln Ile Tyr Gly	
100 105 110	
TAT GAC AGC AGG TTC AGC ATT TTT GGG AAG GAC TTC CTG CTC AAC TAC	548
Tyr Asp Ser Arg Phe Ser Ile Phe Gly Lys Asp Phe Leu Leu Asn Tyr	
115 120 125	
CCT TTC TCA ACA TCA GTG AAG TTA TCC ACG GGC TGC ACC GGC ACC CTG	596
Pro Phe Ser Thr Ser Val Lys Leu Ser Thr Gly Cys Thr Gly Thr Leu	
130 135 140 145	
GTG GCA GAG AAG CAT GTC CTC ACA GCT GCC CAC TGC ATA CAC GAT GGA	644
Val Ala Glu Lys His Val Leu Thr Ala Ala His Cys Ile His Asp Gly	
150 155 160	
AAA ACC TAT GTG AAA GGA ACC CAG AAG CTT CGA GTG GGC TTC CTA AAG	692
Lys Thr Tyr Val Lys Gly Thr Gln Lys Leu Arg Val Gly Phe Leu Lys	
165 170 175	
CCC AAG TTT AAA GAT GGT GGT CGA GGG GCC AAC GAC TCC ACT TCA GCC	740
Pro Lys Phe Lys Asp Gly Gly Arg Gly Ala Asn Asp Ser Thr Ser Ala	
180 185 190	
ATG CCC GAG CAG ATG AAA TTT CAG TGG ATC CGG GTG AAA CGC ACC CAT	788
Met Pro Glu Gln Met Lys Phe Gln Trp Ile Arg Val Lys Arg Thr His	
195 200 205	
GTG CCC AAG GGT TGG ATC AAG GGC AAT GCC AAT GAC ATC GGC ATG GAT	836
Val Pro Lys Gly Trp Ile Lys Gly Asn Ala Asn Asp Ile Gly Met Asp	
210 215 220 225	
TAT GAT TAT GCC CTC CTG GAA CTC AAA AAG CCC CAC AAG AGA AAA TTT	884
Tyr Asp Tyr Ala Leu Leu Glu Leu Lys Lys Pro His Lys Arg Lys Phe	
230 235 240	

53

003060 2985960

12 43

ATG AAG ATT GGG GTG AGC CCT CCT GCT AAG CAG CTG CCA GGG GGC AGA Met Lys Ile Gly Val Ser Pro Pro Ala Lys Gln Leu Pro Gly Gly Arg 245 250 255	932
ATT CAC TTC TCT GGT TAT GAC AAT GAC CGA CCA GGC AAT TTG GTG TAT Ile His Phe Ser Gly Tyr Asp Asn Asp Arg Pro Gly Asn Leu Val Tyr 260 265 270	980
CGC TTC TGT GAC GTC AAA GAC GAG ACC TAT GAC TTG CTC TAC CAG CAA Arg Phe Cys Asp Val Lys Asp Glu Thr Tyr Asp Leu Leu Tyr Gln Gln 275 280 285	1028
TGC GAT GCC CAG CCA GGG GCC AGC GGG TCT GGG GTC TAT GTG AGG ATG Cys Asp Ala Gln Pro Gly Ala Ser Gly Ser Gly Val Tyr Val Arg Met 290 295 300 305	1076
TGG AAG AGA CAG CAG CAG AAG TGG GAG CGA AAA ATT ATT GGC ATT TTT Trp Lys Arg Gln Gln Gln Lys Trp Glu Arg Lys Ile Ile Gly Ile Phe 310 315 320	1124
TCA GGG CAC CAG TGG GTG GAC ATG AAT GGT TCC CCA CAG GAT TTC AAC Ser Gly His Gln Trp Val Asp Met Asn Gly Ser Pro Gln Asp Phe Asn 325 330 335	1172
GTG GCT GTC AGA ATC ACT CCT CTC AAA TAT GCC CAG ATC TGC TAT TGG Val Ala Val Arg Ile Thr Pro Leu Lys Tyr Ala Gln Ile Cys Tyr Trp 340 345 350	1220
ATT AAA GGA AAC TAC CTG GAT TGT AGG GAG GGT GAC ACA GTG TTC CCT Ile Lys Gly Asn Tyr Leu Asp Cys Arg Glu Gly Asp Thr Val Phe Pro 355 360 365	1268
CCT GGC AGC AAT TAAGGTCTTC ATGTTCTTAT TTTAGGAGAG GCCAAATTGT TTTT Pro Gly Ser Asn 370	1325
GTCATTGGCG TGCACACGTG TGTGTGTGTG TGTGTGTGTG TGTAAGGTGT CTTATAATCT	1385
TTTACCTATT TCTTACAATT GCAAGATGAC TGGCTTTACT ATTTGAAAAC TGGTTTGTGT	1445
ATCATATCAT ATATCATTTA AGCAGTTTGA AGGCATACTT TTGCATAGAA ATAAAAAAAA	1505
TACTGATTTG GGGCAATGAG GAATATTTGA CAATTAAGTT AATCTTCACG TTTTTCGAAA	1565
CTTTGATTTT TATTTTCATCT GAACTTGTGTT CAAAGATTTA TATTAAATAT TTGGCATACA	1625
AGAGATATGA AAAAAAAAAA AAAAAAAAAA A	1656

(2) INFORMATION FOR SEQ ID NO:15:

54

008060" 229899150



(2) INFORMATION FOR SEQ ID NO:16:

(A) LENGTH: 1176 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

ATGGCNGGNA	THCCNGGNYT	NYTNTTYTYN	YTNTTYTTYT	TNYTNTGYGC	NGTNGGNCAR	60
GTNWSNCCNT	AYWSNGCNCC	NTGGAARCCN	ACNTGGCCNG	CNTAYMGNYT	NCCNGTNGTN	120
YTNCNCARW	SNACNYTNAA	YYTNGCNAAR	CCNGAYTTYG	GNCGNGARGC	NAARYTNGAR	180
GTNWSNWSNW	SNTGYGGNCC	NCARTGYCAY	AARGGNACNC	CNYTNCCNAC	NTAYGARGAR	240
GCNAARCART	AYYTNWSNTA	YGARACNYTN	TAYGCNAAYG	GNWSNMGNA	NGARACNCAR	300
GTNGGNATHT	AYATHYTNS	NWSNWSNGGN	GAYGGNGCNC	ARCAYMGNGA	YWSNGGNWSN	360
WSNGGNAARW	SNMGNMGNAA	RMGNCARATH	TAYGGNTAYG	AYWSNMGNTT	YWSNATHTTY	420
GGNAARGAYT	TYYTNYTNAA	YTAYCCNTTY	WSNACNWSNG	TNAARYTNWS	NACNGGNTGY	480
ACNGGNACNY	TNGTNGCNGA	RAARCAYGTN	YTACNGCNG	CNCAYTGYAT	HCAYGAYGGN	540
AARACNTAYG	TNAARGGNAC	NCARAARYTN	MNGGTNGGNT	TYYTNAARCC	NAARTTYAAR	600
GAYGGNGGNM	GNGGNGCNAA	YGAYWSNACN	WSNGCNATGC	CNGARCARAT	GAARTTYCAR	660

1856

TGGATHMGNG	TNAARMGNAC	NCAYGTNCCN	AARGGNTGGA	THAARGGNAA	YGCNAAYGAY	720
ATHGGNATGG	AYTAYGAYTA	YGCNYTNYTN	GARYTNAARA	ARCCNCAYAA	RMGNAARTTY	780
ATGAARATHG	GNGTNWSNCC	NCCNGCNAAR	CARYTNCCNG	GNGGNMGNAT	HCAYTTYWSN	840
GGNTAYGAYA	AYGAYMGNCC	NGGNAAYYTN	GTNTAYMGNT	TYTGYGAYGT	NAARGAYGAR	900
ACNTAYGAYY	TNYTNTAYCA	RCARTGYGAY	GCNCARCCNG	GNGCNWSNGG	NWSNGGNGTN	960
TAYGTNMGNA	TGTGGAARMG	NCARCARCAR	AARTGGGARM	GNAARATHAT	HGGNATHTTY	1020
WSNGGNCAYC	ARTGGGTNGA	YATGAAAYGGN	WSNCCNCARG	AYTTYAAAYGT	NGCNGTNMGN	1080
ATHACNCCNY	TNAARTAYGC	NCARATHTGY	TAYTGGATHA	ARGGNAAAYTA	YYTNGAYTGY	1140
MGNGARGGNG	AYACNGTNTT	YCCNCCNGGN	WSNAAY			1176

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1679 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 111...1259
- (D) OTHER INFORMATION:

- (A) NAME/KEY: Signal Sequence
- (B) LOCATION: 111...167
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GAATTCGGCA	CGAGGGGGAG	CCGCGCGCTC	TCTCCCGGCG	CCCACACCTG	TCTGAGCGGC	60
GCAGCGAGCC	GCGGCCCGGG	CGGGCTGCTC	GGCGCGGAAC	AGTGCTCGGC	ATG GCA	116
					Met Ala	

GGG	ATT	CCA	GGG	CTC	CTC	TTC	CTT	CTC	TTC	TTT	CTG	CTC	TGT	GCT	GTT	164
Gly	Ile	Pro	Gly	Leu	Leu	Phe	Leu	Leu	Phe	Phe	Leu	Leu	Cys	Ala	Val	
			-15				-10						-5			

GGG	CAA	GTG	AGC	CCT	TAC	AGT	GCC	CCC	TGG	AAA	CCC	ACT	TGG	CCT	GCA	212
Gly	Gln	Val	Ser	Pro	Tyr	Ser	Ala	Pro	Trp	Lys	Pro	Thr	Trp	Pro	Ala	
1				5					10						15	

57

008060 2298950

2657

TAC CGC CTC CCT GTC GTC TTG CCC CAG TCT ACC CTC AAT TTA GCC AAG	260
Tyr Arg Leu Pro Val Val Leu Pro Gln Ser Thr Leu Asn Leu Ala Lys	
20 25 30	
CCA GAC TTT GGA GCC GAA GCC AAA TTA GAA GTA TCT TCT TCA TGT GGA	308
Pro Asp Phe Gly Ala Glu Ala Lys Leu Glu Val Ser Ser Ser Cys Gly	
35 40 45	
CCC CAG TGT CAT AAG GGA ACT CCA CTG CCC ACT TAC GAA GAG GCC AAG	356
Pro Gln Cys His Lys Gly Thr Pro Leu Pro Thr Tyr Glu Glu Ala Lys	
50 55 60	
CAA TAT CTG TCT TAT GAA ACG CTC TAT GCC AAT GGC AGC CGC ACA GAG	404
Gln Tyr Leu Ser Tyr Glu Thr Leu Tyr Ala Asn Gly Ser Arg Thr Glu	
65 70 75	
ACG CAG GTG GGC ATC TAC ATC CTC AGC AGT AGT GGA GAT GGG GCC CAA	452
Thr Gln Val Gly Ile Tyr Ile Leu Ser Ser Ser Gly Asp Gly Ala Gln	
80 85 90 95	
CAC CGA GAC TCA GGG TCT TCA GGA AAG TCT CGA AGG AAG CGG CAG ATT	500
His Arg Asp Ser Gly Ser Ser Gly Lys Ser Arg Arg Lys Arg Gln Ile	
100 105 110	
TAT GGC TAT GAC AGC AGG TTC AGC ATT TTT GGG AAG GAC TTC CTG CTC	548
Tyr Gly Tyr Asp Ser Arg Phe Ser Ile Phe Gly Lys Asp Phe Leu Leu	
115 120 125	
AAC TAC CCT TTC TCA ACA TCA GTG AAG TTA TCC ACG GGC TGC ACC GGC	596
Asn Tyr Pro Phe Ser Thr Ser Val Lys Leu Ser Thr Gly Cys Thr Gly	
130 135 140	
ACC CTG GTG GCA GAG AAG CAT GTC CTC ACA GCT GCC CAC TGC ATA CAC	644
Thr Leu Val Ala Glu Lys His Val Leu Thr Ala Ala His Cys Ile His	
145 150 155	
GAT GGA AAA ACC TAT GTG AAA GGA ACC CAG AAG CTT CGA GTG GGC TTC	692
Asp Gly Lys Thr Tyr Val Lys Gly Thr Gln Lys Leu Arg Val Gly Phe	
160 165 170 175	
CTA AAG CCC AAG TTT AAA GAT GGT GGT CGA GGG GCC AAC GAC TCC ACT	740
Leu Lys Pro Lys Phe Lys Asp Gly Gly Arg Gly Ala Asn Asp Ser Thr	
180 185 190	

58

009060" 22985960



THE UNIVERSITY OF CHICAGO

59

1859

TTGTCATTGG	CGTGCACACG	TGTGTGTGTG	TGTGTGTGTG	TGTGTAAGGT	GTCTTATAAT	1391
CTTTTACCTA	TTTCTTACAA	TTGCAAGATG	ACTGGCTTTA	CTATTTGAAA	ACTGGTTTGT	1451
GTATCATATC	ATATATCATT	TAAGCAGTTT	GAAGGCATAC	TTTTGCATAG	AAATAAAAAA	1511
AATACTGATT	TGGGGCAATG	AGGAATATTT	GACAATTAAG	TTAATCTTCA	CGTTTTTGCA	1571
AACTTTGATT	TTTATTTTCAT	CTGAACTTGT	TTCAAAGATT	TATATTAAAT	ATTTGGCATA	1631
CAAGAGATAT	GAAAAAAAAA	AAAAAAAAAA	AAAAATTCCT	GCGGCCGC		1679

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 383 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(ix) FEATURE:

(A) NAME/KEY: Signal Sequence

(B) LOCATION: 1...19

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met	Ala	Gly	Ile	Pro	Gly	Leu	Leu	Phe	Leu	Leu	Phe	Phe	Leu	Leu	Cys
				-15				-10							-5
Ala	Val	Gly	Gln	Val	Ser	Pro	Tyr	Ser	Ala	Pro	Trp	Lys	Pro	Thr	Trp
		1					5					10			
Pro	Ala	Tyr	Arg	Leu	Pro	Val	Val	Leu	Pro	Gln	Ser	Thr	Leu	Asn	Leu
	15					20					25				
Ala	Lys	Pro	Asp	Phe	Gly	Ala	Glu	Ala	Lys	Leu	Glu	Val	Ser	Ser	Ser
30					35				40					45	
Cys	Gly	Pro	Gln	Cys	His	Lys	Gly	Thr	Pro	Leu	Pro	Thr	Tyr	Glu	Glu
			50						55					60	
Ala	Lys	Gln	Tyr	Leu	Ser	Tyr	Glu	Thr	Leu	Tyr	Ala	Asn	Gly	Ser	Arg
		65						70					75		
Thr	Glu	Thr	Gln	Val	Gly	Ile	Tyr	Ile	Leu	Ser	Ser	Ser	Gly	Asp	Gly
	80					85						90			
Ala	Gln	His	Arg	Asp	Ser	Gly	Ser	Ser	Gly	Lys	Ser	Arg	Arg	Lys	Arg
	95					100				105					
Gln	Ile	Tyr	Gly	Tyr	Asp	Ser	Arg	Phe	Ser	Ile	Phe	Gly	Lys	Asp	Phe
110					115					120				125	
Leu	Leu	Asn	Tyr	Pro	Phe	Ser	Thr	Ser	Val	Lys	Leu	Ser	Thr	Gly	Cys
				130					135					140	

Leo

008060" 09658677

1860

Thr Gly Thr Leu Val Ala Glu Lys His Val Leu Thr Ala Ala His Cys  
145 150 155  
Ile His Asp Gly Lys Thr Tyr Val Lys Gly Thr Gln Lys Leu Arg Val  
160 165 170  
Gly Phe Leu Lys Pro Lys Phe Lys Asp Gly Gly Arg Gly Ala Asn Asp  
175 180 185  
Ser Thr Ser Ala Met Pro Glu Gln Met Lys Phe Gln Trp Ile Arg Val  
190 195 200 205  
Lys Arg Thr His Val Pro Lys Gly Trp Ile Lys Gly Asn Ala Asn Asp  
210 215 220  
Ile Gly Met Asp Tyr Asp Tyr Ala Leu Leu Glu Leu Lys Lys Pro His  
225 230 235  
Lys Arg Lys Phe Met Lys Ile Gly Val Ser Pro Pro Ala Lys Gln Leu  
240 245 250  
Pro Gly Gly Arg Ile His Phe Ser Gly Tyr Asp Asn Asp Arg Pro Gly  
255 260 265  
Asn Leu Val Tyr Arg Phe Cys Asp Val Lys Asp Glu Thr Tyr Asp Leu  
270 275 280 285  
Leu Tyr Gln Gln Cys Asp Ala Gln Pro Gly Ala Ser Gly Ser Gly Val  
290 295 300  
Tyr Val Arg Met Trp Lys Arg Gln Gln Gln Lys Trp Glu Arg Lys Ile  
305 310 315  
Ile Gly Ile Phe Ser Gly His Gln Trp Val Asp Met Asn Gly Ser Pro  
320 325 330  
Gln Asp Phe Asn Val Ala Val Arg Ile Thr Pro Leu Lys Tyr Ala Gln  
335 340 345  
Ile Cys Tyr Trp Ile Lys Gly Asn Tyr Leu Asp Cys Arg Glu Gly  
350 355 360

003060" 2985960

601